DOCUMENT RESUME

ED 099 127 PS 007 606

AUTHOR Johnson, Stephen M.; And Others
TITLE Evaluation of Family Intervention through Unobtrusive Audio Recordings: Experiences in Bugging Children.
SPONS AGENCY National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE [74]
NOTE 33p.

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS *Audio Equipment; Behavior Change; *Data Collection; *Family Environment; *Intervention; Literature Reviews; Observation; Parent Education; Parent Participation; *Problem Children; Research Methodology; Research Problems; Tape Recordings
IDENTIFIERS *Naturalistic Observation

ABSTRACT

This report presents outcome data on five referrals for child-family intervention in which a new system for obtaining audio-recorded behavioral data was employed as one of several evaluation criteria. The referred child was asked to wear a radio transmitter during most of his waking hours at home during pre-intervention and termination assessments. The transmitter broadcast to a receiver-recording apparatus in the home which could be activated by an interval timer at predetermined "random" times or by parents predetermined "picked" times when the children were said to typically exhibit problematic behavior. During the project, treatment involved an educational program with the parents in which they were trained to keep data on their child's behavior and to institute behavioral programs to change those behaviors. Results are discussed in terms of the frequency of occurrence of child deviant, parent negative, and parent commanding behaviors. In general, the behavioral outcome data verified that from the two other sources (parent attitude reports and parent collected observation data). It was concluded that the audiorecording procedures provide a viable alternative to employing an observer in the home. (Author/SDH)

BEST COPY AVAILABLE SEP. 23 1974

EVALUATION OF FAMILY INTERVENTION THROUGH UNOBTRUSIVE AUDIO RECORDINGS: EXPERIENCES IN BUGGING CHILDREN

U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Stephen M. Johnson, Andrew Christensen and G. Tom Bellamy

University of Oregon

ED 099127

Abstract

This report presents outcome data on five consecutive referrals for child-family intervention in which a new system for obtaining audio recorded behavioral data was employed as one of several evaluation criteria. The referred child was asked to wear a radio transmitter during most of his waking hours at home during pre-intervention and termination assessments. The transmitter broadcast to a receiver-recording apparatus in the home which could be activated by an interval timer at predetermined "random" times or by parents at predetermined "picked" times when the children were said to typically exhibit problematic behavior. As predicted, child deviant, parent negative and parent commanding behaviors were significantly higher at the pretest picked times than at the random times. The reductions at post-test in all three behavior classes were substantial and significant at the picked but not at the random times. Examination of the individual subject data indicated, however, that substantial reductions in at least two of the three dependent variables occurred in three of the five cases at random times. In general, the behavioral outcome data was veridical with that from the two other sources--parent attitude reports and parent collected observation data. It was concluded that these procedures provide a viable alternative to employing an observer in the home while introducing considerable methodological advantages and enabling naturalistic data collection with greater convenience and at less cost.

PS 007606

BEST COPY AVAILABLE

EVALUATION OF FAMILY INTERVENTION THROUGH UNOBTRUSIVE AUDIO RECORDINGS: EXPERIENCES IN BUGGING CHILDREN[1,2]

Stephen M. Johnson, Andrew Christensen and G. Tom Bellamy

University of Oregon

Behaviorists working in applied settings have typically relied on direct observation in the natural environment for the design, modification and evaluation of treatment programs. In general, they have eschewed the more controlled and convenient vehicle of laboratory analogs for clinical outcome research, primarily on the grounds that behavior in such artificial settings may not be representative of behavior in the subjects' natural habitat (e.g., see Fawl, 1963; Kortlandt, 1962; Lobitz & Johnson, 1974; Martin, Johnson, Johansson, & Wahl, 1974; Milgram, 1963). Yet, a rapidly expanding body of methodological literature demonstrates that the representativeness of naturalistic observation data, as typically collected, may also be affected by a variety of confounding influences. As a result, several investigators have pioneered the development of new procedures for direct observation, including the use of audio recording (Bernal, Gibson, Williams, & Pesses, 1971; Johnson & Bolstad, 1974) and radio transmitters (e.g., Purcell & Brady, 1965; Soskin & John, 1963).

The present work involved further development and adaptation of these methods for assessment in the natural home setting. The method employed represents an integration of the use of radio transmitters as first described by Soskin and John (1963) with the use of interval timers to sample behavior at random intervals without subject awareness (as in Bernal, Gibson, Williams, & Pesses, 1971). In the present report, the procedure was employed as one of several methods to evaluate the outcome of a

BEST COPY AVAILABLE

behavior modification parent training program though its potential applications are viewed as extensive.

All of the children involved in the present intervention program were asked to wear a small, wireless transmitter during most of their waking hours at home in the pre- and post-test assessment periods. The transmitter broadcast to a receiver and tape recorder located inconspicuously in the home. The recorder was activated at predetermined intervals by a timer so that neither the child nor his parents knew precisely when their interactions were being recorded ("random times"). In addition, the recorder could be activated by the parents at preselected times which they had identified as most problematic ("picked times").

As previously indicated, this method was developed in the belief that it could at least attenuate many of the data representativeness problems documented in the literature. While a comprehensive review of this literature cannot be given here, a brief cataloging of these problems and the corresponding advantages of this new method of data collection is offered. Of particular importance are problems of subject reactivity, observer bias, lack of generalizability of observer agreement estimates, and lack of representativeness in time sampling procedures. For comprehensive discussions of these problems, the reader is referred to several current reviews on the topic (e.g., Johnson & Bolstad, 1973; Jones, Reid, & Patterson, 1974; Lipinski & Nelson, 1974).

Subject reactivity concerns the possibility that being observed can affect one's behavior. A potential problem in all naturalistic observation, subject reactivity appears to be especially important in the home environment, where a live observer can be a particularly intrusive addition to a

BEST COPY AVAILABLE

typically private and constricted setting. Several aspects of the observation process may affect the observed behavior and lead to unrepresentative data. First, it is possible that expectations held by observers may influence the subjects' behavior. Observer expectancies may be communicated in some way to subjects who then alter their behavior while being observed. Evidence that this expectancy effect can obtain has been provided by Rosenthal and Fode (1963), Rosenthal and Jacobson (1966), and Fode (1965).

More generalized changes in subject behavior can also occur during observation. It has been demonstrated, for example, that subjects may avoid situations in which they can be observed (Bechtel, 1967) or may decrease overall activity level during observation (White, 1972). Of course, if these generalized changes in behavior remain stable, this sort of reactivity, although providing a biased estimate of unobserved behavior, would not necessarily preclude accurate intrasubject measures of treatment effects. Considerably more threatening to the representativeness of behavioral data, however, is the possibility that specific response sets might result from demand characteristics in the observed settings. During home observations prior to family intervention, for example, parents might wish their child to appear deviant in order to insure acceptance for treatment. Similarly, parents might wish to present their child as more normal after treatment in order to express appreciation to therapists and justify termination. That parents can, in fact, manipulate childrens' behavior in response to instructions to do so has been demonstrated in two studies (Johnson & Lobitz, 1974; Lobitz, W. & Johnson, 1974).

BEST COPY AVAILABLE

These investigations demonstrated the troublesome possibility that the effects of observation may interact with demand characteristics accompanying treatment assessment to alter subject behavior. That such interactions can, in fact, occur has been demonstrated in a classroom setting by Kent, Fisher and O'Leary (1974). These investigators found that, relative to periods of inconspicuous observation, children's off-task behavior increased when an observer was present in the classroom prior to treatment, but decreased as a result of observer presence during treatment. Thus, while covert assessment of off-task behavior revealed an increase in off-task behavior during treatment, measures obtained by an observer in the classroom indicated a decrease. The possibility of such differential reactivity during different observational periods is present in much behavioral research.

Several characteristics of the audio recording system used in the present work appear to diminish subject reactivity problems. First, the procedure is clearly less intrusive than sending a live observer into the home and avoids many of the restrictions on behavior of family members necessary with direct observation (e.g., restrictions in movement in the home). Second, the possibility that observers' expectancies may affect subjects' behaviors is avoided entirely. Finally, the mechanical operation of the recording device, which prevents exact knowledge of when data is being collected, should minimize the possibility of generalized changes in behavior or the operation of more confounding, specific response sets. As a result of these factors, data obtained by the audiotape procedure should be more representative of the subjects' unobserved behavior.

Concern for observer bias in behavioral data reflects the possibility that systematic recording errors may occur as a result of the observers'

BEST COPY AVAILABLE

knowledge of the hypotheses under study, or the status of observed subjects. Observer bias represents a threat to data representativeness because of observers' errors, rather than subjects' behavior. That such information can, in fact, result in systematic recording errors is provided by Azrin, Holz, Ulrich, and Goldiamond (1961), Rapp (1966), and Kass and O'Leary (1970). In the present context, knowledge of whether a family observation occurs before or after treatment represents a potential source of observer bias. Since this information is likely to come from the family itself during live home observations, the audio recording procedure used in the present work again seems preferable. There is no necessary correspondence between the order of coding tapes and the actual order of this recording occurrence, and observers can be reassigned as soon as biasing information occurs in a recorded interaction.

The generalizability of observer agreement data is at issue whenever observations are intermittently monitored for agreement. Observer agreement statistics are typically obtained in natural settings by introducing a second, independent observer into the situation and comparing the two records. There is considerable evidence that the resulting measures overestimate agreement at other times (Reid, 1970; Taplin & Reid, 1973; Demaster & Reid, 1972; Romanczyk, Kent, Diament, & O'Leary, 1973). Observer agreement statistics computed when observers were aware of being checked were considerably higher in these studies than when observer agreement was monitored covertly. In most natural settings, however, covert assessment of agreement among live observers is practically impossible. Thus, the present audio recording system again seems preferable, inasmuch as the coding of audio records can be covertly checked for agreement without

BEST COPY AVAILABLE

difficulty.

A final concern with traditional methods of naturalistic observation is the difficulty and expense involved in obtaining an adequate time sample of important behaviors. Behavioral measures which are obtained in limited time periods may be either unrepresentative of overall behavior or insensitive to changes at other times. For example, an adaptation of the behavioral coding system developed by Patterson, Ray, Shaw, and Cobb (1969) has been used in this laboratory to record interactions among family members immediately prior to the evening meal. Although some significant changes in behavior recorded at these times did occur as a result of treatment, the overall rate of deviant behavior was not affected and the results on these observational measures have not paralleled the magnitude of change recorded by a variety of other outcome measures at either termination (Eyberg & Johnson, 1974) or follow-up (Johnson & Christensen, 1974). The times sampled by this home observation system simply may not have been representative enough to reveal actual changes. Although Patterson (1974a, b) has shown significant reductions in deviant behavior after treatment using this method, Ferber, Keeley and Shemberg (1974) have also been unable to do so and have discussed this failure in light of the representativeness problem.

It was expected that the audiotape procedure would provide both a more sensitive and comprehensive outcome measure than direct observation in the home. First, a much larger part of the day could be sampled. The recorder was mechanically activated for a fifteen minute period at breakfast time, before dinner and after dinner each day without the family's awareness ("random times"). Second, provision was made for the parents to activate the recording devices for fifteen minutes daily at a time which they had

BEST COPY AVAILABLE

identified as the most problematic ("picked times").

It was hypothesized that (a) prior to treatment, the rate of child deviant, parent negative and parent commanding behaviors would be greater during the picked times than during the random times; (b) the rate of the three behavior classes would decrease from pretest to post-test during the picked times; and (c) there would be greater decreases in these behaviors during the picked times than during random times. Reductions during the random times were hoped for but not predicted on the basis of prior failures to show changes in overall summaries of child deviance as a function of treatment (Eyberg & Johnson, 1974; Johnson & Christensen, 1974).

Method

Subjects

Five families with children believed to exhibit active behavior problems in the home participated in the treatment program. One family contained two children who were treated in the program. "Active behavior problems" was used to refer to aggressiveness, destructiveness, disobedience, hyperactivity, temper tantrums, or high rate activity with annoyance value. Families were not accepted into this program if the problem child was severely retarded, had experienced severe and documented brain damage or exhibited behavior problems which would ordinarily cause him to be labeled "autistic" or "shizophrenic."

Four families had both parents in the home while one was father-absent. The educational level of parents ranged from 8 to 14 years with the mean of 11.5. Income level for families ranged form $1,800 to $10,000 annually

with a mean of approximately $6,000. The treated children ranged in age from 4 to 10 with a mean age of 6.6.

Treatment Procedures

Six advanced graduate student therapists and one supervising therapist participated in the treatment program[3]. Therapists were formed into permanent two-person teams with both team members typically participating in each treatment session.

The treatment approach was educational in nature and based on the principles of social learning theory and the techniques of behavior modification. Many of the procedures were patterned after those described by Patterson and his associates (e.g., Patterson, Cobb, & Ray, 1973; Patterson, Ray, Shaw, 1968; Patterson & Reid, 1973). Through a series of no more than twelve weekly sessions, parents were trained to keep data on their child's behavior and to institute behavioral programs to change those behaviors. The treatment procedures have been described in greater detail by Eyberg and Johnson (1974).

Measures of Treatment Outcome

There were four measures of treatment outcome: parent observation data, two parent verbal report measures, and audio-tape observational data taken in the home.

Parent observation data. The therapists required both parents to record the to-be-treated child problem behaviors for a one-week baseline period prior to beginning an intervention for them. Parents recorded the frequency and/or duration of the behavior of interest for a specified time each day. Recording time varied from one-half-hour per day for very high rate or situation specific behaviors to the entire day for lower-rate

BEST COPY AVAILABLE

behaviors. During intervention, behavioral recording continued as long as the behavior remained a focus of treatment.

Verbal report measures. A form of the Becker (1960) Bi-Polar Adjective Checklist was employed to obtain parental description of the treated child.[4] This form has previously been employed by Patterson, Cobb, and Ray (1973) to assess parents' change in perception of their children following treatment.

The Therapy Attitude Inventory was constructed specifically for earlier research on the assessment of parents' satisfaction with the process and outcome of this treatment program and with the therapists involved (Eyberg & Johnson, 1974).

Home observations. The audio equipment used for home observations consisted of a transmitter (Edcor PM-1), a receiver (Edcor St-3 sensatuner), a 24-hour timer (Intermatic Model T 1975), two 15-minute time switches (Mark timers), and two reel-to-reel tape recorders. This particular radio package was selected in order to minimize the possibility of accidental pickup of recorded behavior by others. This system employs a relatively little used, low range band typically employed by businesses to communicate with their representatives in the field. Dial scanning by these special users would presumably be minimal. And, the transmitter's maximum range is approximately 200 feet. All of these factors would make accidental pickup by other users extremely unlikely, but all families were informed of this risk.

The transmitter (dimensions: 3-1/4" x 2-3/4" x 3-7/8"; weight, 11 oz.) was worn on the belt of the target child for time blocks when he was inside

BEST COPY AVAILABLE

the home. Time blocks included morning, afternoon, and evening periods for a total of approximately 5 hours a day. The transmitter broadcast continually during these periods.

All other equipment was stored inside a trunk placed in some convenient location in the home (e.g., a closet). The receiver was wired to two tape recorders, one of which was activated at 3 predetermined 15-minute "random times" each weekday. These random times were usually programmed to occur once around breakfast, once in the late afternoon and once in the evening. The target child and parents were aware that some segments of their interaction would be recorded, but they did not know of the duration or exact time of the recordings. The timer was programmed so that it would not activate the random tape on weekends.

The second tape recorder could be activated by the parents operating another 15-minute timer located on the outside of the trunk. Thus, there was a potential for three 15-minute segments of randomly selected audio material and one 15-minute segment of picked material per day. The parents were instructed to operate this "picked tape" once a day during a regularly problematic time or situation. Four families selected bedtime as the most problematic situation and were instructed to turn the tape recorder on at the point in the bedtime sequence at which behavior problems usually occurred. It was emphasized that the same setting event should be used on each occasion. One set of parents was unable to specify a problematic situation, and they were instructed to turn the tape recorder on at a specified time in the afternoon when the child was said to be particularly problematic.

BEST COPY AVAILABLE

The therapists in each case brought the taping apparati to the home for both assessments and introduced the family to its operation. The parents were provided with written as well as oral instructions in the use of this equipment. The instructions emphasized that parents were not to inform their children about the taping at picked times.

In all cases but one, the recording apparatus was kept in the home for five to seven consecutive week days. One family failed to follow the picked time instructions properly during part of the first week of baseline and consequently was required to repeat the assessment for a second week. All random time data was used for this family. Due to fluctuations in family schedules or failures in following instructions, less than the maximum amount of data was collected. The mean number of 15-minute random time segments at each testing period was 9.10 (2 hours 16.5 minutes) with a range of 6-11 for all regular assessments and a high of 16 segments for the one special case at baseline. The mean number of picked time segments was 5.20 (1 hour 18 minutes) per assessment with a range of 4 to 7.

Three means of censorship were provided to protect the families' privacy: (a) The family could activate a censor switch located on the outside of the trunk, disconnecting the receiver for 15 minutes. (b) The family could listen to and erase any part of the tapes prior to their being listened to and coded by observers. (c) The family could listen to and erase any part of the tapes after their coding by observers. Confidentiality of all assessment materials was assured and censorship was not discouraged. None of the families reported using the censor switch and none requested that they listen to or erase any tapes.

BEST COPY AVAILABLE

Observational system. A new observational code, derived in part from the direct observational system devised by Patterson, Ray, Shaw and Cobb (1969), was developed for use with the audio tapes. A slightly shortened form of the present system was used in a recent study by Johnson and Bolstad (1974). The coding system was designed to provide observation data on all of the verbal behavior of the individual wearing the transmitter and all verbal behavior of those who interact with him.

The tapes were coded by two observers trained comprehensively prior to coding these tapes who practiced weekly to maintain coding proficiency. The first observer coded all tapes and was frequently aware of coding tapes representing baseline or termination assessment. Thirty-six of the 142, 15-minute tape segments (or 25%) were randomly selected and coded by a second observer. The second, calibrating observer was unaware of the baseline-termination status of the tapes being coded. Thus, calibrations done on these tapes are representative of the remaining noncalibrated data. To the extent that observers agree on the calibration tapes, an absence of observer bias is indicated.

Observers coded the tapes in 10-second blocks demarcated by clicks recorded on the second channel of each interaction tape. The coding system contains 16 discrete codes which were reduced to four summary statistics for the purposes of this research. These summary codes were child deviant, parent negative, parent commanding and sibling negative behavior. Because this system is relatively new, the computation of observer agreement was rather comprehensive. Observer agreement percent was computed on each code and each summary statistic. To count as an agreement, both observers were

 BEST COPY AVAILABLE

required to see the same event in the same 10-second block. Agreement on nonoccurrence was, of course, not counted. In addition, correlations of the rate per minute scores obtained by each observer in the 36, 15-minute blocks were computed. Those codes which were used less than five times were not subjected to either analysis. A listing of the codes and summary statistics, together with their associated calibrated occurrences and agreement figures is presented in Table 1. The code composition of each summary statistic is also indicated in thistable.

Insert Table 1 About Here

The mean percent agreement for individual code categories was 82% with a range from 70% to 97%. The mean percent agreement for summary statistics was 82% with a range from 77% to 86%. The mean correlation for individual code categories was .94 with all codes but one showing correlations of .92 or better. Summary statistics yielded very high agreement correlations of .97 or higher. For the data analysis involved in this study, the agreement figures on summary statistics are most directly relevant.

Results

Verbal Report Measures

All nine parents involved in this treatment program gave more favorable descriptions of their children at termination than at baseline on the basis of the Becker (1960) Bi-Polar Adjective Checklist. This checklist was scored on the basis of five factors derived by Patterson (as in Patterson, Cobb, & Ray, 1973). Unpublished data from the present laboratory shows that three of these factors (Tense Disposition, Aggression, and Conduct Problems) are highly correlated, and the sum of these three factors has proven to be a useful and valid single index from this measure (Lobitz, G.

BEST COPY AVAILABLE

& Johnson, 1974; Johnson & Christensen, 1974). The baseline and termination scores of all nine parents were compared using a paired observation t-test. In the family with two treated children, only the parents' ratings of the child initially rated as most deviant were included for analysis. The analysis showed that the reliable improvement was statistically significant ($t = 4.83$, $df = 8$, $p < .005$).

The Therapy Attitude Inventory included seven items concerning the parents' ratings of treatment outcome, and three items reflecting parent ratings of the therapists. All items were rated on a scale from 1 (indicating maximum dissatisfaction or deterioration in condition) to 5 (indicating maximum satisfaction or improvement). Parent ratings were averaged in the family with two treated children. For all parents combined, the average rating of those items relating to therapy was 4.37, or between somewhat favorable and very favorable. The mean score for all parents on the questions concerning ratings of the therapists was 4.33. Thus, both verbal report measures indicated favorable outcome for all cases in agreement with previous research involving this treatment program and these verbal report measures (Eyberg & Johnson, 1974; Johnson & Christensen, 1974).

Parent Observation Data

Parents collected behavioral data on each child behavior problem treated. The average number of problems per family was 6.8 with a range from 4 to 10. The outcome results based on parent data were computed on the basis of the percent reduction from baseline observed in the last three weeks of active treatment for the individual behavioral problem. An average percent reduction in child deviant behavior was then computed for each family. The average percent reduction per family was 77% with a range of 58% to 91%.

BEST COPY AVAILABLE

This estimate is also quite consistent with previous research using the same treatment program and observational methods.

Audio Tape Data

The audio tape data for each case on each of the three dependent variables--child deviance, parent negativeness, and parent commands--are presented in Table 2.[5]

Insert Tables 2 and 3 About Here

Four repeated measures, planned comparisons, were performed on each dependent variable reflecting: (a) a comparison of behavior rates at the pretest between picked and random times, (b) a comparison of the pretest and post-test behavior rates at the picked time, (c) a comparison of the pretest and post-test behavior rates at random times, and (d) a test of the interaction between taped time and assessment time. For each variable, significant effects were predicted for comparisons a, b, and d but not for comparison c. The results of these analyses are presented in Table 3.

As predicted, the rates of child deviant behavior were much higher at baseline in the picked times than in the random times. Child deviant behaviors occurred at 1.70 per minute in the baseline picked itmes, but only at .50 per minute at the baseline random times ($t = 5.01$, $df = 8$, $p < .005$). The direction of this difference was the same in every case. Furthermore, the deviant behavior rate per minute was reduced at termination in every case to an average of .50 per minute ($t = 5.98$, $df = 8$, $p < .005$). On the average, the child deviant behavior rates did not change during the random times and the interaction between taped time and assessment time illustrated in Figure 1 was significant ($t = 4.49$, $df = 4$, $p < .01$).

BEST COPY AVAILABLE

Insert Figure 1 about here

Examination of the individual subject data in Table 2 indicates that there was a substantial drop in child deviant behavior during the random time for two cases, essentially no change for two others, and a rather substantial increase in the fifth case.

The same pattern of results is repeated for parent negativeness. Parent negative behaviors were significantly more frequent at the picked than random times at baseline. The parent negativeness decreased reliably and significantly during the picked times, but not during the random times, and the interaction was significant. Examination of the individual subject data in Table 2 reveals substantial reductions in parent negativism in two cases, little change in two others, and a very substantial increase in case number five. The reductions in parent negativism in case numbers two and four in both the picked and random time segments are noteworthy in that reduction in such behavior was an explicit target of the family intervention. Parents in both cases were taught to use time-out procedures and instructed to eliminate all other forms of aversive control such as spanking, yelling, threatening, etc.

The pattern of results was similar for the parent command variable. Parent commands were significantly higher on the average during the baseline picked times than the baseline random times. Parental commands declined reliably and significantly during the picked, but not during the random times, and the interaction was significant. Examination of the individual subject data in Table 2, however, reveals noteworthy reductions in the parent command rate during the random time for four out of the five cases,

BEST COPY AVAILABLE

but a substantial increase in command rate in case number five. The substantial reductions in cases two and four are particularly noteworthy since parental command rate was a discrete target of treatment in these cases.

In summary, all three predictions were confirmed on all three variables. Child deviant, parent negative and parent commanding behaviors were significantly higher at baseline during the picked than random times. All behavior categories demonstrated significant reductions during the picked times but not during the random times and all interactions were significant. The sibling negative variable was not analyzed because siblings were coded in only two cases and were quite inactive when present (see Table 1).

Although the random tape data did not show the consistent declines in child deviant, parent negative, and parent commanding behaviors demonstrated at the picked times, examination of Table 2 indicates that fairly substantial reductions in at least two of these variables was observed for three of the five families (families two, three and four). In the authors' view, the random tape data provide additional evidence for the effectiveness of the treatment program in these three cases. There was a noticeable reduction in the parent command rate in case number one, but all other changes at random times were negligible. The random tape results for case number five are rather discouraging in showing an increase in every dependent variable at the post-test random times. If reliable, this data may reflect the operation of contrast effects as documented in some other behavior modification research (e.g., see Johnson, Bolstad & Lobitz, 1974; Meichenbaum, Bowers & Ross, 1968). In any case, all of these families will continue to have follow-up assessment on all measures and additional treatment will be made available as necessary.

 BEST COPY AVAILABLE

Discussion

Experience with these radio-assisted audio recording procedures has convinced the present investigators of their considerable utility for assessment in the home. The procedures appear to solve many of the representativeness problems as outlined earlier. On logical grounds, the relatively unobtrusive transmitter should yield less subject reactivity and the constant presence of the device during most waking hours should promote more rapid adaptation. In addition to lessening generalized reactivity to being observed, these advantages would seem to eliminate observer communication of expectancies to the observed and lessen the possibility that subjects' responses would reflect only the demand characteristics of the assessment situation.

In many applications, these procedures would enable investigators to keep observers totally uninformed of the hypotheses tested or the circumstances of assessment, and, in cases where this is impossible, the procedures for checking observer agreement can solve the bias problem if observer reliability remains high. Observer agreement checks can be made randomly, as in the present case, by a second, uninformed observer yielding completely generalizable reliability assessments. In this study, there was nearly perfect agreement in observer fluctuations on the three basic summary statistics (correlations = .97-.99). Additionally, and perhaps most important, these procedures can be used to sample behavior at any time. The data presented here would indicate that this provision is particularly useful for treatment evaluation. Although the procedures for defining the "picked" times employed in this study did not eliminate subject awareness, high density "random" taping at specific intervals would presumably have

BEST COPY AVAILABLE

this effect.

The present experience suggests that audio taping yields behavioral data of sufficient density at even random times to be meaningful. Furthermore, observer agreement in coding this data was high by every criteria employed. And, the present experience suggests that naturalistic behavioral data can be collected with these procedures at considerably more convenience to all involved and at less expense than is typically encountered in employing observers in the home.

The tape data results at picked times were highly veridical with parent data and parent attitude reports. The picked tape data showed a mean reduction in deviant behavior of 70% which is comparable to the average 77% reduction in targeted behavior problems as recorded by parents. In the four cases where bedtime problems were targets of intervention, parent data showed an average reduction of 79% in such problems. Although not statistically significant, the random tape results showed substantial reductions in two or more of the dependent variables for three of the five families. The veridicality of this outcome data is particularly encouraging in view of previous research in which behavioral data has failed to substantiate the degree of treatment effectiveness demonstrated by other data sources (Eyberg & Johnson, 1974; Johnson & Christensen, 1974). The present results were completely consistent with the trends observed in this earlier research, however, in that behavioral data showed greater improvement as it was more specifically related to the behavior problems treated. Thus, in this earlier research, parent data on targeted problems showed the greatest improvement, followed by a significant but modest index of improvement based on observer records of those codes related to treated

BEST COPY AVAILABLE

problems followed by nonsignificant change on a general index of observed deviance (Eyberg & Johnson, 1974; Johnson & Christensen, 1974).

The specificity of the taping system is probably responsible for the enhanced agreement among data sources obtained in this study. Since these behavior modification treatment procedures are frequently aimed at very situation and time specific behavior problems, the assessment procedures must possess sufficiently narrow "bandwidth" to yield high "fidelity." Thus, with only 5 subject families, child deviant, parent negative and parent commanding behaviors were significantly higher at the baseline picked times than at the baseline random times. All three variables decreased significantly from baseline to termination at the picked but not at the random times. In every case, the reductions at the picked times were significantly greater than those at the random times.

It must be acknowledged that the picked and random tapes also differed on the dimension of parent awareness of recording. The change in behavior rates could be a function of this awareness and the associated response of parents to demand characteristics. The available data would suggest that parents of generally "deviant" children can behave so as to cause their children to act significantly more deviant than usual but not significantly less deviant (Lobitz, W. & Johnson, 1974). On the average, children in the present study behaved much more deviantly at pretest picked times than they did at the "usual" random times. At the post-test, the mean picked deviant behavior rate was virtually identical to the mean rate at random times. Thus, the data could merely reflect a "fake bad" set at baseline. Several important considerations mitigate against this

BEST COPY AVAILABLE

interpretation. The first comes from consideration of the Walter and Gilmore (1973) study in which a placebo intervention program was employed for families with child problems quite similar to those dealt with here. Although parental expectancies were successfully manipulated by this placebo procedure, observers who employed the presumably more obtrusive method of going into the homes, recorded slight increases in child deviant behavior at the post-test. Furthermore, parental ratings of child symptom occurrence showed no change with placebo treatment. The second consideration derives from a study (Lobitz, G. & Johnson, 1974) in which the pretest deviant behavior rates of referred and normal children were compared with similar behavior rates of comparable groups in which parents had been asked to "fake good," behave normally or "fake bad." The comparisons showed that the deviant behavior means of both referred and normal children corresponded most closely to the "fake good" condition means of the appropriate comparison groups. These results suggest that most families in a home observation situation may respond more to a generalized social desirability response set rather than to any possible demand characteristics associated with treatment. Other results which argue against the demand characteristic hypothesis in the present research include the fact that the parent data was veridical with the tape data in showing high levels of deviant behavior during the picked times prior to treatment and subsequent decreases after treatment. In addition, substantial reductions in child deviant behavior also occurred at random times for two cases and substantial beneficial changes in parent variables occurred for three cases.

In spite of all these considerations, the logical possibility of the

BEST COPY AVAILABLE

demand characteristic hypothesis still exists. As indicated earlier, the system could be adapted for high density random taping at intervals likely to be problematic. Although less efficient than the present procedures, this change would appear to solve any problems introduced by differential awareness and would, as an additional advantage, eliminate the need for parental manipulation of the equipment.

It must be recognized that the present study is purely descriptive in that no control group was employed. Three control group studies have been completed, however, on very similar populations with comparable treatment and assessment procedures including parent report and observational data. The Walter and Gilmore (1973) study involved comparison with an attention placebo group, while the other two studies involved comparisons with no treatment groups (Johnson, Bolstad, & Lobitz, 1974; Wiltz, 1969). The results of all studies have been consistent in showing either no change (Johnson, Bolstad, & Lobitz, 1974) or slight increases in observed deviant behavior in the control or placebo groups (Walter & Gilmore, 1973; Wiltz, 1969). An absence of significant change in parent ratings of the target child was obtained by Johnson, Bolstad, and Lobitz (1974) and Walter and Gilmore (1973) found no change in parent ratings of the child's symptom occurrence with placebo treatment. Thus, although not completely comparable, the available research would tend to indicate that the kinds of changes observed here would not be observed in similar groups exposed to placebo or no-treatment.

BEST COPY AVAILABLE

References

Azrin, N. H., Holz, W., Ulrich, R., & Goldiamond, I. The control of the content of conversation through reinforcement. Journal of the Experimental Analysis of Behavior, 1961, 4, 25-30.

Bechtel, R. B. The study of man: Human Movement and architecture. Trans-Action, 1967, 4 (6), 53-56.

Becker, W. C. The relationship of factors in parental ratings of self and each other to the behavior of kindergarden children as rated by mothers, fathers, and teachers. Journal of Consulting Psychology, 1960, 24, 507-527.

Bernal, M. E., Gibson, D. M., Williams, D. E., & Pesses, D. I. A device for recording automatic audio tape recording. Journal of Applied Behavior Analysis, 1971, 4 (2), 151-156.

Demaster, B. & Reid, J. B. The efficacy of the stop-check procedure in maintaining the reliability of data collected by observers in quasi-natural settings: Two pilot studies. Oregon Research Institute Research Bulletin, 1972, 12 (8).

Eyberg, S. M. & Johnson, S. J. Multiple assessment of behavior modification with families: Effects of contingency contracting and order of treated problems. Journal of Consulting and Clinical Psychology, 1974, in press.

Fawl, C. Disturbances experienced by children in their natural habitats. In R. G. Barker (Ed.), The stream of behavior. New York: Appleton-Century-Crofts, 1963. Pp. 99-126.

Ferber, H., Keeley, S. M., & Shemberg, K. M. Training parents in behavior modification: Outcome of and problems encountered in a program after Patterson's work. Behavior Therapy, 1974, 5, 415-419.

BEST COPY AVAILABLE

Fode, K. L. The effect of experimenters' and subjects' anxiety and social desirability on experimenter outcome bias. Unpublished doctoral dissertation, University of North Dakota, 1965.

Johnson, S. M. & Bolstad, O. D. Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. A. Hamerlynck, L. C. Handy, and E. J. Mash (Eds.), Behavior change: Methodology, concepts and practice. Champaign, Illinois. Research Press, 1973. Pp. 7-67.

Johnson, S. M. & Bolstad, O. D. Reactivity to home observation: A comparison of audio recorded behavior with observers present or absent. Journal of Applied Behavior Analysis, 1974, in press.

Johnson, S. M., Bolstad, O. D., & Lobitz, G. K. Generalization and contrast phenomena in behavior modification with children. In L. A. Hamerlynck, L. C. Handy, and E. J. Mash (Ed.), Parenting: Directions, change and maintenance of healthy family behavior. 1974, in press.

Johnson, S. M. & Christensen, A. Multiple criteria follow-up of behavior modification with families. Unpublished manuscript, University of Oregon, Eugene, 1974.

Johnson, S. M. & Lobitz, G. K. Parental manipulation of child behavior in home observations. Journal of Applied Behavior Analysis, 1974, 7, 23-31

Jones, R. A., Reid, J. B., & Patterson, G. R. Naturalistic observations in clinical assessment. In P. McReynolds (Ed.), Advances in psychological assessment. Volume 3. San Francisco, California: Jossey-Bass, 1974, in press.

BEST COPY AVAILABLE

Kass, R. E. & O'Leary, K. D. The effects of observer bias in field-experimental settings. Paper presented at a symposium entitled "Behavior Analysis in Education," University of Kansas, Lawrence, April 1970.

Kent, R. N. Fisher, J. B., & O'Leary, K. D. Observer presence as an influence on teacher and child behavior in a classroom setting. Unpublished manuscript, State University of New York, Stony Brook, 1974.

Kortlandt, A. Chimpanzees in the wild. Scientific American, 1962, 206, 128-138.

Lipinski, D. & Nelson, R. Problems in the use of naturalistic observation as a means of behavioral assessment. Behavior Therapy, 1974, 5, 341-351.

Lobitz, G. K. & Johnson, S. M. Normal vs. deviant children: A multimethod comparison. Paper presented at the meeting of the Western Psychological Association, San Francisco, California, April 1974.

Lobitz, W. C. & Johnson, S. M. Parental manipulation of the behavior of normal and deviant children. Paper presented at the meeting of the Western Psychological Association, San Francisco, California, April 1970.

Martin, S., Johnson, S. M., Johansson, S., & Wahl, G. The comparability of behavioral data in laboratory and natural settings. In L. A. Hamerlynck, L. C. Handy, and E. J. Mash (Eds.), Parenting: Directions, change and maintenance of healthy family behavior. 1974, in press.

Meichenbaum, D. H., Bowers, K. S., & Ross, R. R. Modification of classroom behavior of institutionalized female adolescent offenders. Behavior Research and Therapy, 1968, 6, 343-353.

Milgram, S. Behavioral study of obedience. Journal of Abnormal and Social Psychology, 1963, 67, 371-378.

BEST COPY AVAILABLE

Patterson, G. R. Intervention for boys with conduct problems: Multiple settings, treatments and criteria. Journal of Consulting and Clinical Psychology, 1974, in press. (a)

Patterson, G. R. Retraining of aggressive boys by their parents: Review of recent literature and follow-up evaluation. In F. Lowy (Ed.), Symposium on the seriously disturbed preschool child. Canadian Psychiatric Association Journal, 1974, 19, 142. (b)

Patterson, G. R., Cobb, J. A., & Ray, R. S. A social engineering technology for retraining the families of aggressive boys. In H. E. Adams and I. P. Unikel (Eds.), Issues and trends in behavior therapy. Springfield, Illinois: C. C. Thomas, 1973. Pp. 139-224.

Patterson, G. R., Ray, R. S., & Shaw, D. A. Direct intervention in families of deviant children. Oregon Research Institute Research Bulletin, 1968, 8 (9).

Patterson, G. R., Ray, R. S., Shaw, D. A., & Cobb, J. A. Manual for coding family interactions, sixth revision, 1969. Available from ASIS National Auxiliary Publications Service, in care of CCM Information Service, Inc., 909 Third Avenue, New York, New York 10022. Document #01234.

Patterson, G. R. & Reid, J. B. Intervention for families of aggressive boys: A replication study. Behavior Research and Therapy Journal, 1973, 11, 383-394.

Purcell, K. & Brady, K. Adaptation to the invasion of privacy: Monitoring behavior with a miniature radio transmitter. Merrill-Palmer Quarterly, 1965, 12, 242-254.

Rapp, D. W. Detection of observer bias in the written record. Cited in R. Rosenthal, Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

Reid, J. B. Reliability assessment of observation data: A possible methodological problem. Child Development, 1970, 41, 1143-1150.

BEST COPY AVAILABLE

Romanczyk, R. G., Kent, R. W., Diament, C., & O'Leary, K. D. Measuring the reliability of observational data: A reactive process. Journal of Applied Behavior Analysis, 1973, 6 (1), 175-184.

Rosenthal, R. & Fode, K. L. The effect of experimenter bias on the performance of the albino rat. Behavior Science, 1963, 8, 183-189.

Rosenthal, R. & Jacobson, L. Teacher's expectancies: Determinants of pupils IQ gains. Psychological Reports, 1966, 19, 115-118.

Soskin, W. F. & John, V. P. The study of spontaneous talk. In R. G. Barker (Ed.), The stream of behavior. New York: Appleton-Century-Crofts, 1963. Pp. 228-281.

Taplin, P. S. & Reid, J. B. Effects of instructional set and experimental influence on observer reliability. Child Development, 1973, 44, 547-554.

Walter, H. & Gilmore, S. K. Placebo versus social learning affects in parent training procedures designed to alter the behaviors of aggressive boys. Behavior Therapy, 1973, 3, 361-377.

White, G. D. Effects of observer presence on mother and child behavior. Unpublished doctoral dissertation, University of Oregon, Eugene, 1972.

Wiltz, N. A. Modification of behaviors of deviant boys through parent participation in a group technique. Unpublished doctoral dissertation, University of Oregon, Eugene, 1969.

BEST COPY AVAILABLE

Footnotes

[1]This research was supported by National Institute of Mental Health Grants MH 19633 and MH 25298. The authors would like to express their appreciation to Russell Glasgow and James Whinston who assisted in the completion of various aspects of this work.

[2]Reprints may be obtained from Stephen M. Johnson, Psychology Clinic, 1679 Agate, University of Oregon, Eugene, Oregon 97403.

[3]Therapists were Manuel Barrera, G. Tom Bellamy, Andrew Christensen, Katy O'Banion, Elizabeth Steinbock and Marie Rering Witt. The first author served as supervisor.

[4]All assessment instruments used in this research are available from the authors on request.

[5]Case number 3 involved a family in which a 9- and 10-year-old boy were both involved in the treatment program, and each boy wore the transmitter on alternate days. This was the only family in which siblings could not be reliably discriminated in coding. As a result, all child and parent behaviors were coded only in those segments when both children were together and all data was averaged. These procedures resulted in the loss of only 1.33 minutes of data in pretest picked time and 12 minutes in the post-test picked time.

BEST COPY AVAILABLE

Table 1

Observer Agreement Computed by Percent and Correlation

For Each Code and Summary Statistic

Code	Total Calibrated Occurrence	Percent Agreement	Correlation
Talk	2143	93	.99
* Positive Command	193	81	.98
* Negative Command	39	79	.59
* Suggestive Command	17	76	.95
* Repeated Command	131	87	.99
* Future Command	0	-	-
** Disapproval	3	-	-
'** Threat	30	73	.92
'** Whine	179	84	.96
'** Yell	219	80	.97
'** Verbal Negativism	82	70	.96
'** Tease	1	-	-
'** Physical Negativeness	1	-	-
'** Disapproving Tone	164	73	.93
' Cry	101	97	.99
' Demand Attention	82	78	.99
Commands	379	86	.99
Parental Negativeness	239	77	.97
Sibling Negativeness	6	83	1.00
Child Deviant	622	83	.99

* Codes included in the command category.
** Codes included in the parent or sibling negative category.
' Codes included in the child deviant category.

BEST COPY AVAILABLE

Table 2

Rate per Minute of Child Deviant, Parent Negative and Parent Commanding Behavior at Pretest and Post-Test at Random and Picked Times

	Picked Times			Random Times		
	Pretest	Post-Test	% Change	Pretest	Post-Test	% Change
Child Deviant						
Case						
1	.683	.197	-71%	.454	.515	+13%
2	1.823	.883	-52%	.835	.711	-15%
3	1.480	.333	-78%	.291	.073	-75%
4	2.680	.686	-74%	.606	.345	-43%
5	1.824	.413	-77%	.314	.773	+146%
Mean	1.670	.502	-70%	.500	.483	-
Parent Negative						
Case						
1	.300	.016	-95%	.218	.187	-14%
2	.101	.050	-50%	.181	.070	-61%
3	1.120	.222	-80%	.028	.027	- 4%
4	2.206	.372	-83%	.598	.142	-76%
5	.392	.227	-42%	.036	.288	+700%
Mean	.824	.177	-70%	.212	.143	-
Parent Commands						
Case						
1	.450	.082	-82%	.370	.269	-27%
2	.253	.250	- 1%	.441	.187	-58%
3	1.356	.500	-63%	.080	.036	-55%
4	1.340	.220	-84%	.697	.233	-67%
5	.757	.440	-42%	.153	.439	+187%
Mean	.831	.298	-54%	.348	.233	-

BEST COPY AVAILABLE

Table 3

Planned Comparison t-tests on Child Deviant, Parent Negative and Parent Commanding Behavior[a]

Comparison	Variable		
	Child Deviance	Parent Negative	Parent Command
Picked vs. Random Times at Pretest	5.01****	2.63**	2.48**
Pretest vs. Post-Test at Picked Times	5.98****	2.61**	3.20***
Pretest vs. Post-Test at Random Times	.08	.30	.69
Interaction: Taped Time X Assessment Time[aa]	4.49***	2.28*	2.21*

*$p < .05$
**$p < .025$
***$p < .01$
****$p < .005$

[a]All tests are one-tailed ($df = 8$ for first three comparisons; $df = 4$ for Interaction).

[aa]The interaction for child deviance is illustrated in Figure 1.

BEST COPY AVAILABLE

Figure 1

Mean Child Deviant Behavior at Pretest and Post-Test
During Random and Picked Times